Assessments of Simulated Performance of Alternative Architectures for Command and Control: The Role of Coordination

Susan P. Hocevar [1]
Department of Systems Management
555 Dyer Road
Naval Postgraduate School
Monterey, CA 93943
(831)656-2249
fax: (831)656-3407
Shocevar@nps.navy.mil

William G. Kemple, David Kleinman, Gary Porter
C4I Academic Group
Naval Postgraduate School
589 Dyer Road
Monterey, CA 93943

Abstract

This paper presents the results of the fourth in a sequence of experiments conducted by the Adaptive Architectures for Command and Control research team. The focus of this study is on the relative effectiveness of three organizational structures in the conduct of a simulated Joint Task Force mission. Two of the three organizational architectures were optimized, using pre-experimental modeling, to limit the amount of inter-nodal coordination. These two structures varied in level of workload (4-node vs. 6-node). The third structure was based on a more traditional, functional design that required more inter-nodal coordination than them model-based structures. Effectiveness was evaluated in terms of performance on the more predictable primary mission tasks as well as some less predictable tasks and a measure of general protection of the force. Overall, there is limited evidence that the 6-node structure designed to reduce inter-nodal coordination performed more effectively than the other two on the primary mission tasks. There is also limited evidence that the traditional structure that required more coordination in accomplishing primary tasks, was more effective than the model-based structures in responding to the less predictable tasks. This evidence supports the value of coordination capabilities in responding to situations of uncertainty.

1. Background

The issues of change and continuity in command and control, the theme for the 1999 CCRP Symposium, are at the heart of the research conducted by the Adaptive Architectures for Command and Control (A2C2) research team. This team comprised of university, government, industry and military participants has conducted, over several years, a series of simulation experiments that examine varying dimensions of command and control (structural distribution of

[1] This work was supported by both the Office of Naval Research, Cognitive and Neural Sciences Division, A2C2 and the Institute for Joint Warfare Analysis (IJWA) at the Naval Postgraduate School.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **1999**	2. REPORT TYPE	3. DATES COVERED **00-00-1999 to 00-00-1999**

4. TITLE AND SUBTITLE **Assessments of Simulated Performance of Alternative Architectures for Command and Control: The Role of Coordination**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Postgraduate School,555 Dyer Road,Monterey,CA,93943**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
1999 Command & Control Research & Technology Symposium

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**21**	

assets, authority, communications systems, decision making, etc.) and their relationship to organizational performance and organizational adaptability. The purpose of this paper is to present some of the results of the most recent experiment conducted in 1998 (Experiment 4). This experiment extends the findings of prior experiments that have been reported in previous symposium proceedings.

A major focus of the previous (third) experiment was on the willingness of officers, playing the role of a Joint Task Force (JTF), to change the organization's structure in the face of a trigger event. The hypothesis, based both on change theory and previous A2C2 research, was that participants would choose a structure that was more familiar (A0-6) over an alternative structure (A1-4) that had been determined through pre-experimental modeling to be optimal for accomplishing the mission tasks facing the JTF. Full reports on this experiment are presented in the 1998 conference proceedings [see for example, Benson, et al. 1998; Curry et al. 1998; Entin et al. 1998; Handley, et al. 1998; Hocevar, 1998; Hocevar et al. 1998]. In Experiment 3, teams were all initially trained in structure A0-6 (traditional). After the training, teams were briefed that there was going to be a significant reduction in assets (due to a conflict in another region). Three reduced asset structures were proposed. One structure maintained the key structural form of A0-6 with assets reduced. The other two structures (A1-4 and A1-5[2]) were derived from pre-experimental modeling to reduce the amount coordination required across organizational "nodes." The pre-experimental modeling hypothesized that structures that reduced the need for coordination by assigning task-defined assets to individual commanders would perform more effectively than the more traditional/functionally defined structure that would necessitate internodal coordination for task accomplishment.

Two important findings in Experiment 3 influenced the design and focus of Experiment 4. First, all of the nine six-person teams of military officers (rank 03-05) participating in this experiment chose the more traditional and coordination intensive structure (A0-6) as the one with which they could most effectively perform the mission rather than either of the two "optimized" alternatives. This, on the surface, confirmed the hypothesis of preference for "proximal" over "optimal" structure. However, interesting insights emerged from the officers' justification for this choice [Hocevar, 1998; Hocevar *et al.*, 1998]. They explained that A0-6 would allow them to be more effective both due to it's similarity with traditional military structure and their experience with this structure in the training phase of the simulation. This finding reflects the military's strong emphasis on training and rehearsal. The second major rationale was the more insightful. Teams also argued that, because the A0-6 structure required coordination among component commanders, it offered performance advantages. Specifically, they saw the required coordination as both facilitating adaptive response to unexpected events and encouraging commanders to maintain a full mission perspective and not become overly focused on their assigned tasks. Some officers argued that this structure would thus provide better overall protection of the force.

A second, unanticipated, finding from Experiment 3 was that the A0-6 structure chosen by all 9 teams outperformed the A1-4 structure that had been predicted to be more effective based on the pre-experimental optimization and modeling. Benson *et al*. (1998) provided some analyses to explain this anomalous finding. Because A1-4 was designed for nodal commanders to perform

[2] A1-4 and A1-5 represented 4-node and 5-node structures that allowed for an assessment of the impact of workload.

tasks autonomously, they frequently had to simultaneously utilize multiple assets toward a significant target. In the training preparation for the experiment, participants had practiced coordination, but this practice had emphasized coordination with other team members. The reduced accuracy of teams using the A1-4 structure suggests that their performance may be an artifact of training that did not adequately prepare players to perform multi-asset tasks autonomously. A second possible explanation is the confounding aspect of workload. While A1-4 reduced the workload inherent in inter-nodal coordination, teams in this structure still faced the same mission tasks as A0-6, but with two fewer nodal commanders.

2. Relevant Theoretical Concepts[3]

A major focus of this study is on the role of coordination capability in adapting to environmental uncertainty or less routine task demands. Contingency theorists argue that organizational effectiveness is influenced by the degree of "fit" between the requirements of the environment and the characteristics of the organization [Burton & Obel, 1998]. For example, substantial research in organization theory has demonstrated that a task-based (divisional) organization with largely autonomous units can be highly effective in situations where the environment can be segmented (e.g., by task domains) and is fairly stable and predictable [Duncan, 1979]. This type of situation has limited information processing needs and low requirements for coordination.

However, organization theory also suggests that this type of structure may be less effective at responding to unanticipated requirements for coordination across divisions [Mintzberg, 1993]. In situations of uncertainty, lateral coordination capabilities can enhance an organization's ability to adapt to meet the increased information processing requirements [Galbraith, 1977, 1995]. While the definition of "uncertainty" in organization theory has not been totally uniform, the following characteristics cited by Burton and Obel [1998] are common: complexity (i.e., number of variables), unpredictability (reduced awareness of relevant data), interdependence, and rate of change.

Weick and Roberts [1993] developed the concept of "collective mind" to explain the effective behavior of highly interdependent activities in the high risk environment of aircraft carriers. They found that success in such a dynamic, complex, fast-paced environment results from "heedful interrelating" of activities that creates a pattern of joint action. Using a very different sized organization, Waller [1999] examined the effectiveness of flight crews in response to nonroutine events. She found that crews that were able to more quickly identify a non-routine event, distribute critical information, re-prioritize their tasks, and adjust task distributions were more effective than those whose behaviors were more fixed in the dominant routine activities.

3. Purpose

The focus of this study (Experiment 4) is to clarify the findings from Experiment 3 and further examine the role of coordination in performance of both planned and unanticipated tasks. The first issue is to clarify the role of training adequacy on the performance of the A1-4 structure. The experimental design and simulation training for Experiment 4 was modified to assure that participants were fully familiar with their role, the assets they were responsible for, the overall

[3] For a more detailed discussion of theoretical concepts see [Hocevar, *et al.* 1998]

structure they were part of, and the software requirements for "coordinated attacks." With the biasing factor of training differences removed, the optimized, model-based Architectures are predicted to outperform the traditional structure that has increased requirements for inter-nodal coordination. A second clarification of Experiment 3 is to understand the role of workload in performance. The design of the alternative organizational structures tested in Experiment 4 (described below) address this question.

Finally, a focal of purpose of this paper is to examine the hypothesis that emerged from the justification given by the military officers for their choice of A0-6 structure in Experiment 3 [Hocevar, 1998; Hocevar *et al.*, 1998]. As noted above, many argued that this structure would be more effective in response to unanticipated events and more effective in the general protection of the force because it required coordination. They felt the practiced coordination would make them more adaptive and keep individual nodal commanders focused on the total mission rather than autonomous sub-tasks. Thus, this study examines the following research question: When faced with the need to respond to an unanticipated, complex, task, does a structure that requires some inter-unit coordination provide a performance advantage over a structure that minimizes coordination by using a task-based design?

4. Method

4.1 *Sample*

Study participants were 54 military officers in varying Master of Science programs at the Naval Postgraduate School. Table 1 presents the breakdown by service. Eleven of the participants were foreign national officers from a variety of nations (e.g., Norway, Turkey, Ukraine, Hungary, Senegal). Of these, six were navy, four army and one air force. More than 90% of the total group were of rank 03 and 04. Four of the foreign national officers were rank 05 and one was 02. Of the U.S. officers, 13 were Special Operations. All but one of the participants were male.

Table 1. Service Representation of Study Participants

USN	USA	USMC	USAF	USCG	International	TOTAL
19	12	6	4	2	11	54

Teams of six members were constituted by random assignment, maintaining a balance across teams in terms of service, operational vs. support experience, and U.S. vs. foreign national origin. At least one foreign national officer was assigned to each team. While not part of the experimental design, this provided a unique opportunity to simulate not just a joint service, but a coalition engagement.

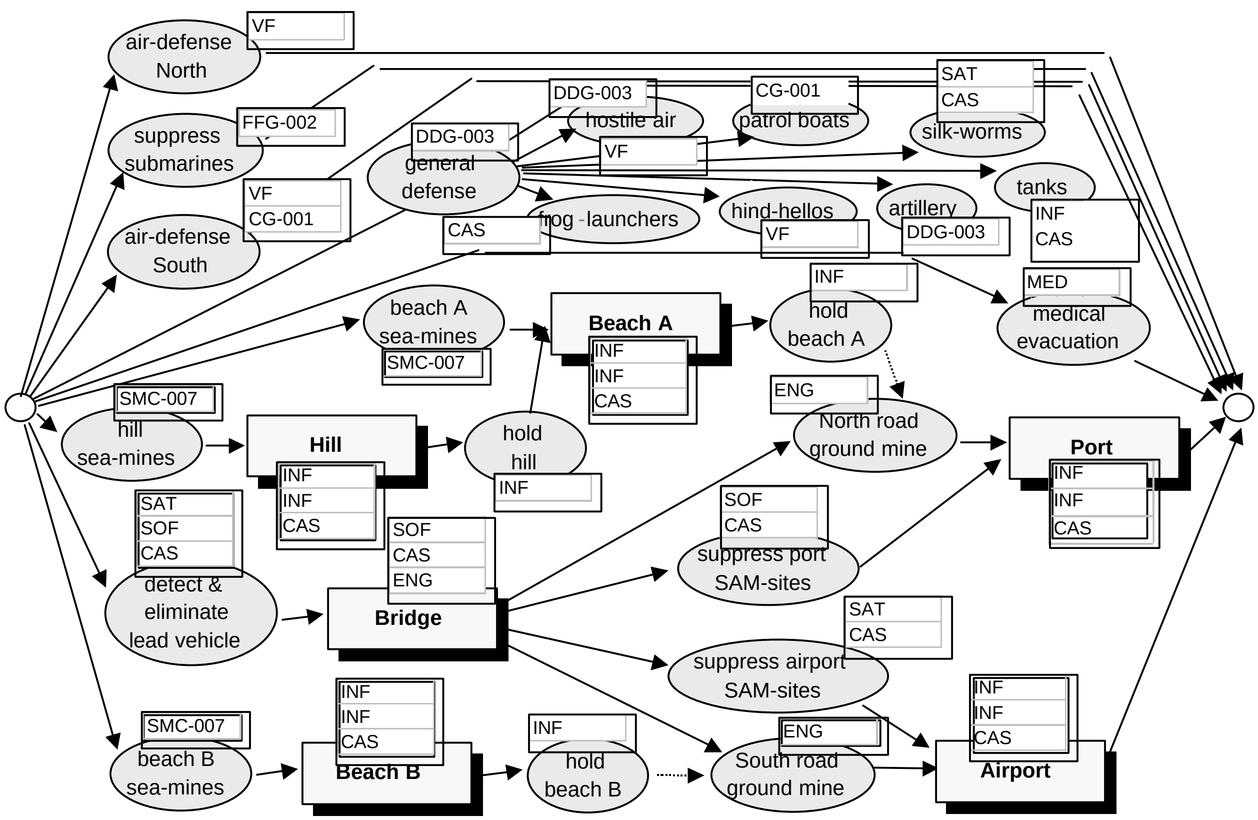
4.2 *Scenario*

The teams were asked to play the role of a Joint Task Force (JTF) and presented a scenario in which the U.S. is taking action in support of an ally, Country Green, that has been invaded by

neighboring Country Orange. The ultimate objectives of this mission are to secure Country Green's Airport and Port. A mission briefing document that outlined a specific chronology of mission tasks to be undertaken by the JTF was distributed to all participants. A greatly simplified version is listed below. A graphic summary is presented in Figure 1:

1. Amphibious forces will land and take North and South Beach after clearing mines. (Note: 2 INF (Infantry) and 1 CAS (Close Air Support) are shown in the figure as one option for accomplishing these mission tasks)
2. Prior to taking N. Beach, infantry (INF) and air support will seize and hold the hill overlooking the beach.

Figure 1. Task-Resource Graph



3. Infantry will move down roads from S. Beach toward airport and from N. Beach to Port clearing mines and enemy tanks.
4. Special Operations Force (SOF) and satellite (SAT) must determine which of two roads the enemy plans to use for insertion of forces by assessing traffic. Once the enemy "lead vehicle" is identified, it should be destroyed as well as the bridge being used by that vehicle, while retaining second bridge for friendly traffic.
5. Armored counterattack forces are believed to be at the Airport and Port. If present, the must be identified and destroyed.

6. Both the Port and Airport must be captured and held. The attack on the Airport has priority and should occur first if they cannot be attacked simultaneously.

The above mission tasks are assumed to represent a high level of certainty. They all MUST occur, and the sequencing, location, and assets required for each task are clearly specified and documented for team participants. With the exception of clearing the land and sea mines, all of the above mission tasks are somewhat complex in that they required multiple assets to accomplish (e.g., infantry (INF) and close air support (CAS)). Because of the focus in the paper on the role of coordination, the major mission tasks requiring multiple assets will be one focus of the analysis.

The mission briefing document also defined "Other Tasks That Can Occur Throughout The Operation." Of interest to this research paper are only those that require multiple assets. The two tasks in this category are detecting and destroying Silkworm and SAM sites. These tasks represent higher level of uncertainty in that they "pop up" as the simulation unfolds. In contrast with the major mission tasks, the exact location and timing of these tasks is not known by the JTF decision makers. These tasks are also of moderate complexity in that two different assets (close air support with either satellite or SOF) are required.

4.3 *Independent variables: Organizational structure*

As in Experiment 3, three organizational structures were evaluated in terms of their impact on performance effectiveness. Figure 2 presents the three structures and shows the distribution of assets among the JTF decision makers.[4] A0-6 represents the more traditional, functionally defined structure and is identical to the one used in Experiment 3. A multi-objective optimization procedure similar to the one described by Levchuk *et al.* [1998] was used to define the A1-4 and A1-6 structures. This optimization process is based on a complex analysis of resource and task vectors. The major objectives used in the optimization process were reduced requirements for coordination, geographic proximity, and workload. The primary differentiating feature between the two model-based structures is number of "nodes." Thus, the impact of workload on performance can be assessed by comparing A1-4 and A1-6. Comparison of A0-6 with the two model-based structures provides a test for this study's focal question of the impact of coordination requirements on organizational performance.

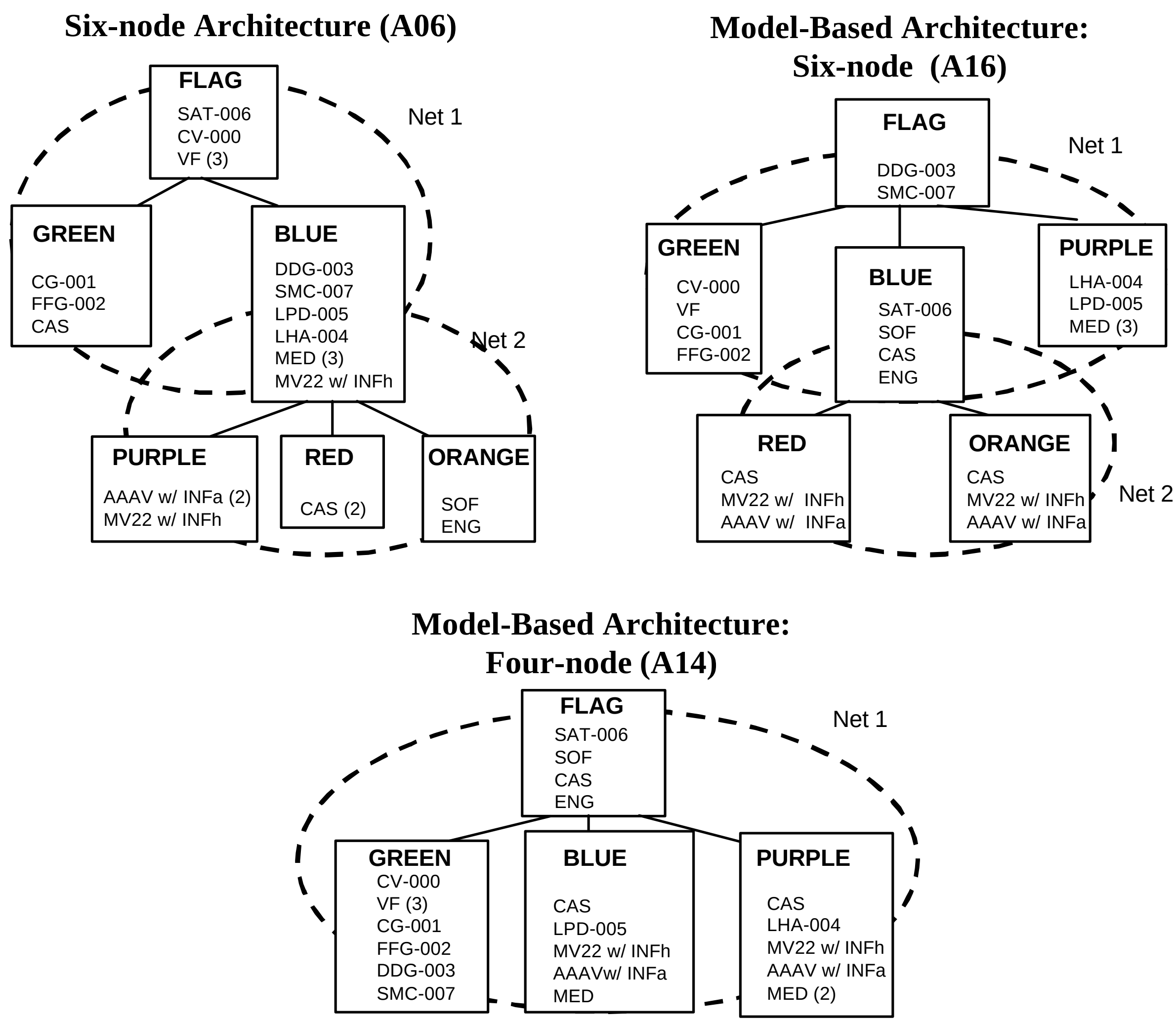
Figure 2 also indicates the use of communication nets as part of the structural definition of each architecture. For both A1-6 and A0-6, communication was divided between two networks with "Blue" playing the communication coordinator. A1-4 represented a "flattened" organization with reduced hierarchy and a single communication net.

The experiment was designed such that each team would employ two different organizational structures in the accomplishment of the specified mission using the DDD-III simulator. Assignments were counterbalanced across the total set for order effect. In the first round of trials, 3 teams were assigned structure A1-4. Given teams defined in sets of six, there were

---

[4] Because of refinements in the modeling, A1-4 in Experiment 4 is not exactly equivalent to that in Experiment 3, thus data from the two experiments could not be combined.

Figure 2. Three Organizational Architectures



enough participants to constitute an additional team. The same was true (of a different set of 3 teams) in round two. These teams of four were balanced in terms of demographics as described above. Thus, the overall design generated data trials from seven teams for A1-6 and A0-6 and data trials from six teams for A1-4.

The key characteristics of organizational structure to be examined as independent variables in this research are: optimized vs. non-optimized (i.e., low vs. high coordination requirements for major mission tasks) and number of organizational nodes (i.e., workload).

4.4 *Training*

In addition to the mission briefing document summarized above, participants received additional information to review in preparation for the experiment. Two matrices described both the enemy

and friendly assets and the capabilities of each of these assets. A third matrix defined each specific task and the resource requirements needed for its accomplishment. This matrix included recommended and optional “force packages” that could be used (e.g., 1 INF, 1 CAS, and DDG to take N. Beach). Finally, participants also received a training manual in how to use the Distributed Dynamic Decisionmaking-III (DDD-III) simulation software.[5]

The question addressed in Experiment 3 of the role of “familiarity” over “optimality” as factors influencing choice of structure was not part of Experiment 4. Thus “familiarity” did not need to be a manipulated variable. This allowed participants to receive equivalent training in both of the organizational structures they would use in the simulation trials. Training occurred in two phases. First, participants received 1-2 hours of individual level training in the use of the DDD software. (The additional training was provided to increase criterion proficiency with the software as determined by both self-selection and observation of trainers). At the end of the software training, individuals were given a handout that graphically presented the organizational structure (similar to Figure 2) they would use in their first trial.

Each team was scheduled into the simulation lab at for two 3-hour blocks that combined two training and one data trial for each of the two structures they were assigned. When teams arrived at the lab, they self-determined who would be assigned to each specific node command position. This enhanced the “fit” between expertise and the type of tasks and assets each would be responsible for in the mission simulation. During the first two hours, teams went through two training runs that were similar to the ultimate mission simulation in terms of the major task requirements and chronology, though somewhat slower in pace and with less demand in terms of hostile threats requiring defensive action. The third hour was dedicated to the data trial and follow-up paper-and-pencil self-report measures [see Entin, 1999] and a brief After Action Review (AAR).

4.5 *Dependent Measures*

Entin [1999] has prepared a companion paper on Experiment 4 that is presented in this *Proceedings*. This paper provides additional data on the comparative effectiveness of the three organizational architectures. The primary measures used by Entin are derived from two sources. He utilizes the aggregate measures of performance derived from DDD-III: Mission Score and Strength Score. He also evaluates the task-level and aggregate ratings of expert observers of both task effectiveness and teamwork behaviors.

To evaluate the training adequacy, the analysis of Mission and Strength scores will also be presented here. However, the major analytic focus of the paper presented here is on more “micro” measures of effectiveness generated by DDD-III and reported in the software’s “Dependent Variable File” for each run of the simulation. These measures include objective assessments of task-level performance. “Task Accuracy” is calculated based on the extent to which teams met the vector of requirements (i.e., used an appropriate combination of assets) in accomplishing that task. Scores are based on 100% and decrements reflect both inadequate force application or when combined attacks are not coordinated within a specified time limit [see Kleinman & Young, 1997 for more details on task score calculations].

[5] Readers interested in details on the DDD-III simulation software should see Kleinman *et al.* [1996].

There are two categories of Task Accuracy measures used in this analysis. The first is the set of seven major offensive mission tasks that were defined in the mission briefing. These tasks represent a high degree of certainty in that the requirements, timing, asset options, location, etc. are all known in advance; thus they are taken as measures of highly *predictable* tasks. Defining these tasks as "certain" does not suggest that the environment they were to take place in was totally predictable. In particular, the defend tasks against aircraft, submarines, patrol boats were known in general, but the particular location and timing was uncertain. Most of the defensive tasks required only a single asset and thus could not be used to test the role of coordination in adapting to an uncertain environment. However, two task categories were identified as representing both higher uncertainty and the requirement of coordination[6] -- hostile SAM and Silkworm missile sites. These two task categories required teams to be "heedful" in order to "observe" their appearance in the mission domain [following Weick & Roberts, 1993]. They also required an interruption of the priority offensive mission tasks and change in the immediate task focus of certain decision makers [following Waller, 1999]. The two categories of Task Accuracy measures are listed below:

Tasks Representing Conditions of Certainty	Tasks Representing Conditions of Uncertainty
North Beach	Silkworms
South Beach	SAMs
Hill	
Airport	
Lead Vehicle	
Bridge	
Port	

The above discussion has implications to the assessment of the focal concept of "adaptability." In Experiment 3, a trigger event (reducing assets) was used as the situational change requiring teams to decide how to adapt by identifying their "preferred" organizational structure. In the absence of a similar significant trigger event in Experiment 4, "adaptation" could not be operationalized in the same way. In this study, "adaptation" is defined in terms of a team's ability to shift focus from the primary "certain" mission tasks to the more uncertain, unanticipated events that involve some coordinated resource response. This characterization of adaptation is in line with the tenets of organization theory described in the literature review. Thus, the test of "adaptability" will assessed by the comparative performance of the three organizational architectures on the "uncertain" tasks defined above.

In addition to the task accuracy scores, the DDD-III dependent variable file records the number of organizational nodes involved in each task. These data will be analyzed to determine the degree coordinated behavior was used and the correlation of degree of coordination with performance results. The final measure to be analyzed from the DDD-III dependent variable file is the "number of hostile penetrations." This score is used to test the officers' rationale for the choice of A0-6 in Experiment 3: Is a structure requiring coordination among organizational

[6] Dave Kleinman, designer of the DDD-III software simulation, played a key role in the identification of tasks appropriate to the experimental conditions of interest. In Figure 1, "Tanks" also appear as a defend task that requires two assets. They were not considered to be as "uncertain" as the two missile defense tasks because of the way they were spawned by contact with infantry, and the consequent certainty of awareness, location, and response.

nodes (i.e., A0-6) more effective in the general protection of the forces than those that are task-designed to increase nodal autonomy (A1-6 and A1-4).

5. Results and Discussion

As described in the Method section above, the experimental design allowed for six data trials for A1-4 and seven data trials for A1-6 and A0-6. Unfortunately, the simulator "crashed" during one of the A1-4 data trials thus decreasing the sample for this cell to five. In analyzing data from training runs, only the second (final) training runs were included. Again, one trial run "crashed" for a team conducting A0-6. In addition, while DDD-III task level data were available, the summary Mission and Strength scores were lost from two additional training runs (A0-6 and A1-4). These variations from the original design will be reflected in the samples sizes reported in analyses below.

5.1 *Training Effect*

The first set of analyses focused on the question from Experiment 3 as to whether a difference in training effectiveness biased the comparative performance of the three Architectures. Two overall measures of performance derived from DDD-III were used. Mission score reflects the extent to which teams accomplished all tasks (not just the 7 primary tasks listed above) with the appropriate resources. At the beginning of a simulation, the score starts at zero and cumulates to a possible score of 100%. Strength score starts at 100% and decrements accrue to reflect such things as enemy "hits," collisions, and mission tasks attempted with inadequate resources (as defined by the vector of task requirements).

An initial test of the difference between the last training run and the trial run on Mission and Strength scores shows that the teams were still making adjustments to their strategy in the trial. The Mission score significantly increased in the trial run as compared with training; but Strength score significantly decreased (see Table 2). This finding suggests that teams were focused on

Table 2. Comparison of Mission and Strength Scores for Training vs. Trial Runs

	Training (N=17)	Trial (N=19)
Mission	72.2 (10.1)	81.5* (12.9)
Strength	85.0 (3.9)	80.1* (6.7)

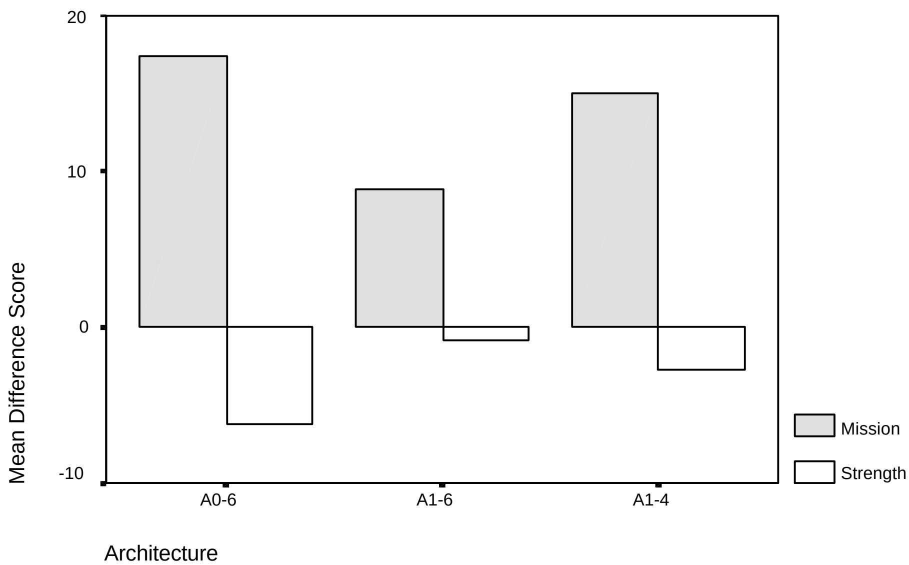
* $\underline{p}<.05$ using 2-tailed t-test (df=34)

improving their performance on the mission tasks and this priority led to a small, but significant decrease in their effectiveness in dealing with factors such as hostile unanticipated threats.

To determine whether there was a different effect for training depending on the type of organizational architecture, a follow-up evaluation was conducted. A within-team difference score was calculated between each team's Mission score on the second (final) training run and

trial run for a specific architecture. A similar difference score was calculated for Strength. A one-way ANOVA found no significant differences for the effect of architecture on the training vs. trial difference scores for either the Mission or Strength scores. Thus, there is no bias based on training differentially effecting the performance of the three architectures. Figure 3 illustrates that the pattern of increase in Mission scores and decrease in Strength scores between training and trial runs is consistent across the three architectures.

Figure 3. Comparison of Performance: Training vs. Trial



5.2 ***Comparative Performance on Primary Mission Tasks by Three Architectures***

As described in the Method section above, each team had three opportunities to apply a given architecture to the accomplishment of the simulated mission. Two of these were training trials and one was the actual "data" trial. However, DDD-III data were being gathered during the second training trial as well as the final data trial. Because of the limited sample size in each condition (i.e., Architecture type), data from the final (second) training run were combined with

the data from the actual intended trial run to attempt to increase the statistical power of analyses. There are two justifications for this. First, the trends in the results for the final training runs followed a similar pattern to those from the trial runs. This was demonstrated by a multivariate ANOVA using Mission and Strength as the dependent variables and Architecture, Order, and the interaction of these two factors as independent variables. The interaction term was not significant ($p<.10$) thus demonstrating that the relationship between Architecture and performance measures was not dependent on whether data were from the second training run or the actual data trial.

The second justification for including the training data is conceptual rather than statistical. As stated in the introduction, one of the major objectives of the A2C2 research project is to evaluate how "adaptive" an architecture is. The experimental assessment of "adaptation" is measured in

this study by how well teams respond to less predictable tasks. But, another part of adaptation can be interpreted by how well an architecture performs, even when teams are fairly inexperienced with that architecture. It is argued here, that by including the final training data along with the trial data, we can better assess the effectiveness of the three architectures in regard to the varying research questions.

5.2.1 *Performance as Measured by Overall Mission and Strength Scores*

Before presenting the task-level data that are the focus of this study, the results of the overall measures provided by DDD-III are useful as a type of performance baseline.[7] An ANOVA comparing the three Architectures on both Mission and Strength scores was conducted and the results are presented in Table 3. The results show that the F-statistic was not significant for

Table 3. Comparison of Architectures on Mission and Strength

	A0-6 (N=12)	A1-6 (N=14)	A1-4 (N=10)	F-test (df=2,33)
Mission	73.08 (11.54)	81.71 (8.74)	75.5 (16.37)	n.s.
Strength	82.0 (5.62)	85.71 (3.81)	78.40 (6.75)	$p<.01$

Note: Cells represent mean and (std dev).

Mission, thus the mean scores cannot be interpreted as different for the three Architectures. In contrast, the Strength scores are different, with a *post hoc* LSD test showing that the performance of A1-6 is significantly stronger than that of A0-6 and A1-4 ($p<.05$). Interestingly, the pattern of means for Mission score is the same (i.e., A1-6 outperforming A0-6 and A1-4), but the high error variance does not allow these differences to reach the level of statistical significance. Thus, there is only partial support for the hypothesis that the model-based structure (A1-6) outperforms the traditional (A0-6) structure. The finding that A1-6 also outperforms the model-based A1-4 architecture suggests that the higher workload required by A1-4 may impede performance that is reflected in the Strength score.

5.2.2 *Performance as Measured by Accuracy on Primary Offensive Mission Tasks*

One of the intended contributions of this study was to examine the utility of task-level data derived from DDD-III to enhance our understanding of the relative performance of varying organizational architectures. To that end, the task accuracy score for each of the seven major mission objectives (defined as high in predictability) were analyzed using ANOVA with Architecture as the independent variable. Table 4 shows the ANOVA results. Included in this table are a composite score for these seven tasks (Total Accuracy) and the previously presented overall Mission score. [For purposes of clarity, standard deviations are omitted from this table.]

---

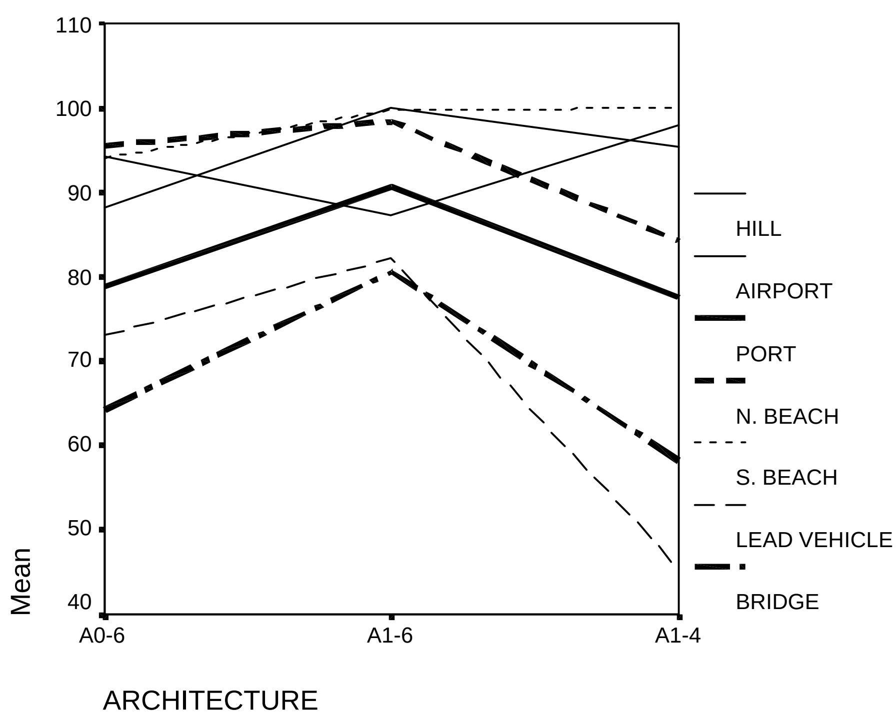
[7] As noted in the Method section, a complementary analysis of teams' performance in this experiment, with more elaboration on these and expert observer data can be found in Entin [1999].

Table 4. Mission Task Accuracy by Architecture				
Task	**A1-6** (N=14)	**A0-6** (N=13)	**A1-4** (N=11)	**F sig. level**
N. Beach	98.4	95.6	84.3	0.18
S. Beach	99.8	94.1	100	0.28
Hill	87.4	94.2	98	0.32
Airport	100	88.2	95.5	0.22
Lead Veh.	82.1	73.1	45.5	0.09*
Bridge	80.6	64.3	58.2	0.41
Seaport	90.6	78.8	77.6	0.56
MISSION score	81.7	73.1	75.5	0.19
TOTAL Accuracy	91.3	84	79.9	.08*

* Significant F-statistic, p<.10
Accuracy values range from 0 to 100%

Clearly, there is limited statistical evidence for the impact of Architecture on task performance for the primary offensive mission tasks. Only the Lead Vehicle task and the composite Total Accuracy measure show significant main effect for Architecture ($p < .10$). *Post hoc* tests to compare pairs of means show the model-based structure A1-6 significantly outperforming both the more traditional structure A0-6 and the second model-based structure A1-4 (using LSD test, $p<.05$). While not reaching statistical significance, Table 4 shows a similar pattern of performance for all but one of the task measures (A1-6 outperforming A0-6 and A1-4). This pattern is presented graphically in Figure 4.

Figure 4. Task Accuracy for Primary Mission Tasks by Architecture



Even though these results include both training and trial data, the sample size is still quite small and thus the statistical analysis has limited power. However, both the table of means and graphic presentation above show a consistent pattern. To further examine this pattern, a more descriptive presentation of the results was compiled based on a within-team comparison. Given the experimental design, each team generated both training and trial data for two different structures. A simple tally was made of the number of times one architecture outperformed the other (for the same team) for each of the 7 primary mission tasks. For example, if Team A was assigned to both A1-6 and A0-6, the task accuracy scores for both the training and trial events for "Hill" were coded in one of three categories: A1-6 > A0-6; A0-6 > A1-6; or no difference (i.e., scores for Hill accuracy under both architecture conditions were equal). This approach disregards the degree of difference in accuracy scores. However, it can be argued that a lower accuracy score (even if small in absolute value) represents an objective determination of lesser performance in either damage caused to the enemy or greater incurred losses to friendly forces.

Table 5 presents a summary of the tabulations described above. By way of example, the first data cell shows that there were a total of 42 trials (3 x 7 x 2 = 42; where 3= # of teams assigned to both A1-4 and A1-6; 7 = # of mission task accuracy scores evaluated; and 2 = # of data runs (final training and trial)). Of these trials, 24 showed the team performing equally well using both organizational architectures. However, when they performed differently (18 of the 42 events) 83% of these differences showed A1-6 performing with more accuracy than A1-4.

Table 5 Within Team Comparisons of Mission Task Accuracies by Architectures

	# of tasks	% of total tasks	% of task differences
A1-4 vs. A1-6			
A1-6 > A1-4	15	36%	83%
A1-4 > A1-6	3	7%	17%
no difference	24	57%	
A1-4 vs. A0-6			
A1-4 > A0-6	9	26%	50%
A0-6 > A1-4	9	26%	50%
no difference	17	48%	
A1-6 vs. A0-6			
A1-6 > A0-6	13	37%	68%
A0-6 > A1-6	6	17%	32%
no difference	16	46%	

The pattern of results comparing A1-6 with A0-6 shows a similar trend. While there were fewer trials (due to one "crashed" run and no available data), on 68% of the tasks where there were differences in performance, A1-6 had higher performance accuracy than A0-6. Interestingly, the comparison of A1-4 and A0-6 shows no overall advantage of one over the other. This presentation of the data, while not demonstrating statistical significance, suggests a pattern of results that merits further study to increase the sample size and thus the power of inferential

analyses. The summary shows evidence of support for the hypothesis that the model-based structure (A1-6) with reduced coordination requirements outperforms the more traditional A0-6 with higher coordination requirements. Because A1-4 and A1-6 were both derived from pre-experimental optimization modeling, the difference in overall performance is likely the result of differences in workload.

It is important to note that the tasks analyzed in this section were those characterized as fairly high certainty (see previous discussion in Method). In addition to providing limited support for the hypothesis mentioned above, the results also reflect some tenets on organization design derived from organization theory. Specifically, when environments can be segmented (e.g., by tasks) and those tasks are fairly predictable, a divisional organization with fairly autonomous, task-defined units limited coordination requirements can be quite effective. This aligns with the finding that A1-6 has the strongest task accuracy performance on the primary mission tasks. The question to be addressed next is the relative effectiveness of the three Architectures in responding to tasks of greater uncertainty and unpredictability.

5.3 *Performance on Less Predictable, "Adaptive" Tasks*

As described in the Method section, there were only two task categories that met the definition of higher uncertainty to allow for comparison with the results presented above. These were the measures of accuracy for action against enemy SAM and Silkworm missile sites. These tasks appeared several times in the simulation. They would "pop up" at times and places that were somewhat unpredictable. There were 3 possible Silkworm threats in the training runs, and 5 in the trial runs. Similarly, there were 2 possible SAM threats in the training and 3 in the trial runs. It should be noted that Silkworms and SAMs were not of equal unpredictability (as defined by degree of awareness of relevant data [Burton & Obel, 1998]). Some of the SAM threats were "prerequisite" tasks for accomplishment of the primary mission tasks of Port and Airport. Because of this, there were DDD-III based messages sent when a team would attempt to attack either of these primary mission sites without first addressing the pre-requisite SAM site(s). In contrast, Silkworm sites were totally dependent on the teams' being "heedful" of the more unpredictable task requirements and determining when and if it was appropriate to shift task focus from the primary tasks in order to eliminate this threat.

Table 6 presents the ANOVA results testing the effect of Architecture on team performance on these less predictable tasks. As with the analysis of primary mission tasks, the differences in mean performance do not reach the level of statistical significance for either of the missile tasks. However, for the Silkworm tasks, the F-statistic is approaching a level of statistical significance ($\underline{p}$=.13), suggesting that more data to increase statistical power should be gathered in future experiments[8]. It is noteworthy that the pattern of mean accuracy scores in Table 6 is different from the consistent pattern shown for the more predictable primary mission tasks in Table 5. For these more unpredictable tasks, the pattern shows A0-6 as outperforming A1-6 and A1-4. If these preliminary trends can be supported with further data, this would confirm the hypothesis derived from Galbraith [1977] and others who propose that coordination capability improves

[8] Due to the way accuracy scores for the two missile tasks are compiled in the DDD-III dependent variable file, it was not possible to do a more descriptive presentation comparable to Table 5.

organizational effectiveness with uncertain environments and less predictable, complex and interdependent tasks.

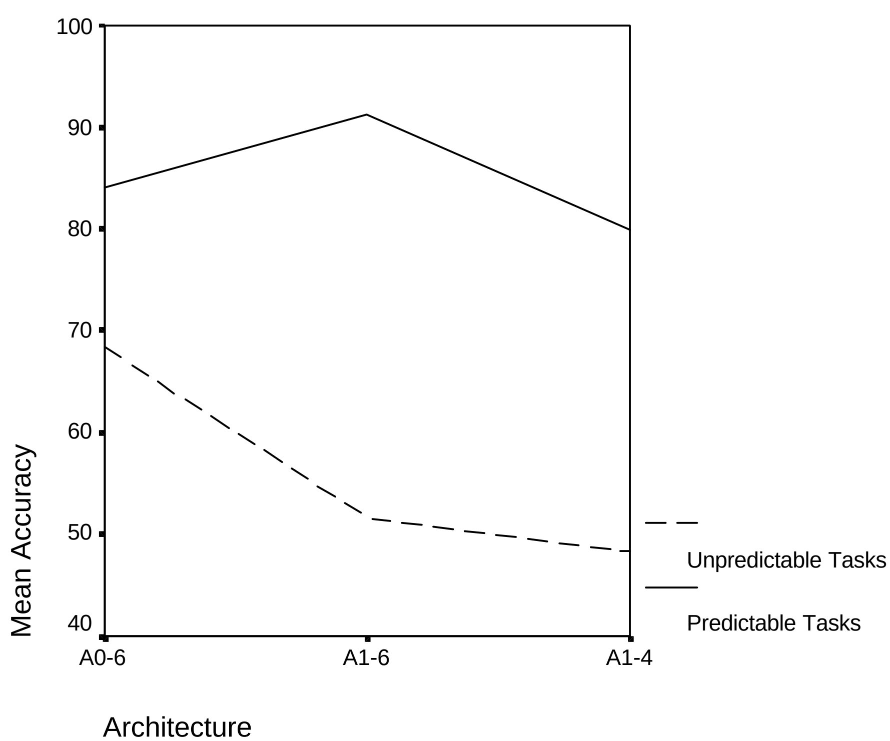
Table 6. Comparative Performance on Less Predictable, "Adaptive" Tasks

	A1-6 (N=14)	**A0-6** (N=13)	**A1-4** (N=11)	**F-sig. level**
Silkworms	219	289	188	0.13
SAMs	116	156	125	0.48
# enemy penetrations	6.1	6.2	11.3	.006*

* Significant F-statistic, p<.01
Silk and SAM performance: (%Accuracy) x (# accomplished);

Figure 5 is a graphical presentation of the different overall pattern of mean task accuracy for two composite scores. “Predictable Tasks” represents the Total Accuracy measure reported in Table 4 that reflects the average accuracy score for the 7 primary mission tasks. “Unpredictable tasks” is an average of the overall accuracy scores for the two sets of missile tasks. This average was adjusted by the total number of missile task opportunities so that the two graphic measures would be comparably based on a scale of 0 to 100%. This pattern of results suggests it is

Figure 5. Accuracy scores for Predictable and Unpredictable Tasks by Architecture



important to assess the differential performance of structures on tasks that vary in terms of uncertainty and the role of coordination as it affects performance.

5.3.1 *Effect of Architecture on "General Protection of the Force"*

One aspect of Table 6 has yet to be discussed. As noted in the introduction to this paper, several of the military officers who participated in Experiment 3 argued that one advantage of structure A0-6 was that the coordination requirements would contribute to greater effectiveness in attending to the general protection of the force. They expressed concern that the optimized structures designed to increase sub-unit autonomy, while more efficient, ran the risk of losing a total mission focus [Hocevar *et al.*, 1998]. This concern is often given as a limitation for divisional organizations [Mintzberg, 1993].

One measure gathered by DDD-III is the number of hostile penetrations. This represents the number of times in a simulation that the enemy is able to successfully attack a friendly asset by using aircraft, sea craft or missiles. For all of these sources of threat, detection, identification, timing, and location are somewhat unpredictable. Only the effectiveness of responses to missile threats have been analyzed so far because the other "threat" tasks require only a single asset to respond and thus do not require coordination. However, the overall number of hostile penetrations provides a good indicator of the team's ability to protect against these less predictable events and thus a measure for "general protection of the force." It also demonstrates the extent to which teams are "heedful" of less predictable events [Weick & Roberts, 1993] and can effectively redirect assets away from the more programmed offensive tasks, to the less routine defensive tasks [Waller, 1999].

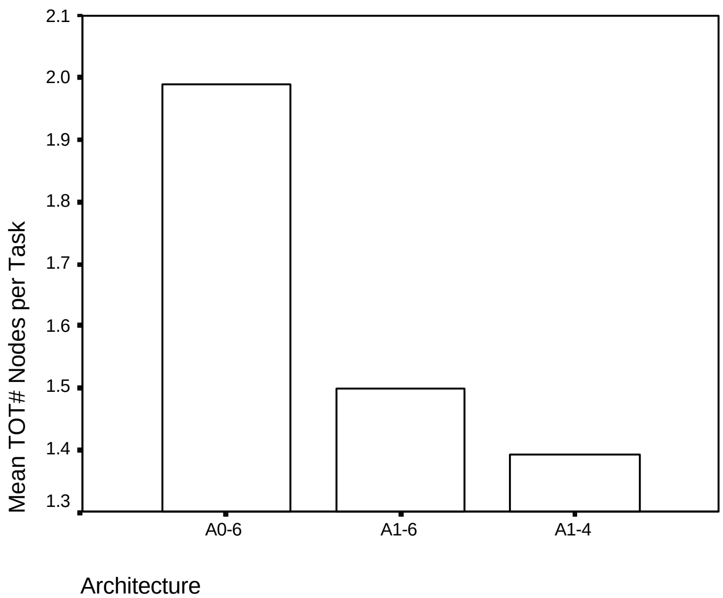
As shown in Table 5, the ANOVA found a significant effect for Architecture on the average number of hostile penetrations ($p<.01$). *Post hoc* tests of difference between means (using LSD test) showed A1-4 to be significantly less effective than either A0-6 or A1-6. This result does not allow for a clear interpretation as to whether the major differentiating factor is coordination capability or workload. It is possible that coordination capability (highest in A0-6, second highest in A1-6) led to the higher performance on this measure. Alternatively, the greater workload experienced by teams in A1-4 could be the cause. Further experiments will be needed to clarify this question. Independent of the resolution of this question, the finding suggests an important caution. The results presented on the effect of Architectures on the primary mission tasks showed A1-4 achieving relatively equivalent performance with A0-6, even though it was reduced in manning by 1/3. The findings presented here on the "defensive" capabilities of A1-4, both as reflected in the missile tasks and the number of enemy penetrations, suggest that it is important to include a broad range of measures of effectiveness when determining the impact of reduced manning.

5.4 *The Role of Coordination in Performance*

While not definitive, the evidence in the section above suggests that coordination plays a role in effective performance. This concluding section of the Results examines this question in more detail. At the simplest level, it is important to demonstrate that the three structures met the objective of pre-experimental modeling in terms of differentiation by degree of coordination

required. Figure 6 shows the means plot derived from an ANOVA that found a significant effect ($p<.001$) for Architecture on the average number of "nodes" used per task (using the 7 primary mission tasks). As we would expect from the fact that A1-6 and A1-4 were based on pre-experimental modeling to reduce the amount of required coordination, the means for these two structures are significantly lower than that for A0-6 ($p<.001$, using LSD test).

Figure 6. Average Number of Nodes per Task by Architecture



Of greater interest is whether the amount of coordination used in the execution of tasks correlates with performance as measured by task accuracy.[9] Pearson correlation coefficients were calculated between the number of nodes per task and consequent task accuracy. Table 7 shows the correlational results for the two separate missile tasks, classified as higher in uncertainty, and the composite Task Accuracy score representing the more predictable primary mission tasks.

Table 7. Correlation between Number of Nodes Involved and Task Accuracy

	Task		
	Silkworms	SAMs	Total Task Accuracy
r	.71**	.92**	.23 (ns)

r = Pearson correlation coefficient; 2-tailed test:
** $p < .01$
N=27; Note: Only two six-node structures (A0-6; A1-6) included

Because the A1-4 structure was the most constrained in the likelihood for coordination (due to only 4 nodes), the table shows the results for only A0-6 and A1-6.[10] The correlations between amount of coordination and performance outcome for both of the more uncertain missile tasks

---

9 Hutchins *et al.* [1999], in this *Proceedings,* present a complementary analysis of qualitative communication behaviors as indicators of coordination processes and their relationship with performance effectiveness.
10 A similar pattern emerged from the analysis of the total sample, with the same statistical significance.

are highly significant ($\underline{p}$<.01). For the more certain tasks represented by the Total Task Accuracy score, the correlation is not significant. These results provide additional evidence in support of Galbraith's [1977, 1995] theory that increased information processing capability through lateral coordination mechanisms enhances performance in situations of higher uncertainty.

6 Conclusions

There is clear evidence reported in this study that the pre-experimental modeling to reduce coordination requirements for A1-4 and A1-6 was accomplished. Having established this distinction, it was possible to test the difference this independent variable had on task performance. Two categories of tasks were evaluated. For the primary mission tasks, characterized as low in uncertainty, A1-6 shows some evidence of outperforming A0-6 and A1-4. This suggests that reducing the requirements for coordination can be a factor in improving organizational performance when tasks are routine, though complex.

Results that tentatively support a different conclusion were presented for the missile tasks that represented higher uncertainty. Performance on these tasks as well as in the "general protection of the force" as measured by number of enemy penetrations, showed A0-6, with the highest requirements for coordination, performing as well or better than A1-6 and A1-4. In addition, the correlational results provide further evidence for the value of coordination capability in the effective response to uncertainty. While some of the findings reported here were not statistically significant, the patterns are consistent. In addition, these patterns are supported by theories of organization design that propose the value of coordination capabilities as an adaptive mechanism for responding to environmental and task uncertainty [e.g., Galbraith, 1977, 1995; Duncan, 1979; Burton & Obel, 1998]. For example, Weick's and Roberts's [1993] work on aircraft carriers suggests that situations that are fluid, high risk, and highly interdependent require teams to be "heedful" of threatening events and practiced in creating a joint response to these events. The results presented here similarly suggest that an organizational design that has some degree of practiced coordination capability may be more effective in responding to unpredictable, high risk situations such as the missile tasks and the threats requiring defensive action in the A2C2 scenario. This tentative conclusion is also supported by the recent work of Waller [1999] with flight crews in simulator training. She found that a significant predictor of effectiveness was the extent to which teams could more quickly change task priorities and re-distribute tasks in response to a crisis. Practiced coordination may facilitate that adaptive response.

An obvious limitation of this study is the limited sample size and resulting constraints on statistical power. It will be important to follow-up this experiment with both computer simulations and human-in-the-loop experiments with larger samples to further test these key concepts. The major design criteria for Experiment 4 was to replicate and resolve questions raised in Experiment 3. For this reason, it was not possible to significantly change the task domain of the scenario used in the simulation. Future experiments should increase the number of tasks that represent the characteristics of uncertainty. This will allow a more rigorous test of the role of coordination mechanisms in adaptability. With a larger number of such tasks, it will also be possible to qualitatively examine [following Waller, 1999] the processes teams use in dealing with nonroutine events. There is evidence reported here supporting the performance benefits

that can be gained by an efficiency-oriented, task-based design with limited coordination requirements. There is also evidence for the benefits of an organization design that requires and thus reinforces coordination. These two characteristics of organization design are not necessarily mutually exclusive. The challenge is to determine the appropriate balance.

7. References

[Benson *et al*., 1998] Benson, R. Kemple, W.G., Kleinman, D., Porter, G., Serfaty, D. An example of model-based empirical research: A soup-to-nuts evaluation of alternative C2 architectures. Proceedings for the 1998 Command & Control Research and Technology Symposium (pp. 34-43). Monterey, CA, June 1998.

[Burton & Obel, 1998] Burton, R.M. & Obel, B. *Strategic Organizational Diagnosis and Design: Developing Theory for Application (2nd Ed.)*. Boston: Kluwer Academic Publishers, 1998.

[Curry *et al.*, 1998] Curry, M.L., Kleinman, D.L., & Pattipati, K.R. On mission planning: A foundation for organizational design. *Proceedings for the 1998 Command & Control Research and Technology Symposium* (pp. 1-10). Monterey, CA, June 1998.

[Duncan, 1979] Duncan, R. What is the right organization structure? Decision tree analysis provides the answer. *Organization Dynamics*, Winter: 59-79, 1979.

[Entin, 1999] Entin, E. Optimized Command and Control Architectures for Improved Process and Performance. Paper presented at the 1999 Command and Control Research and Technology Symposium, Newport RI, June 1999.

[Entin *et al.*, 1998] Entin, E., Serfaty, D. & Kerrigan, C. Choice and performance under three command and control architectures. *Proceedings for the 1998 Command & Control Research and Technology Symposium* (pp. 132-137). Monterey, CA, June 1998.

[Galbraith, 1995] Galbraith, J.R. *Designing organizations*. San Francisco: Jossey Bass, 1995.

[Galbraith, 1977] Galbraith, J.R. *Organization design*. Reading, MA: Addison-Wesley, 1977.

[Handley *et al.*, 1998] Handley, H.A.H., Zaidi, Z.R. & Levis, A.H. Pre-experimental modeling of adaptive organizational architectures. *Proceedings for the 1998 Command & Control Research and Technology Symposium* (pp. 44-53). Monterey, CA, June 1998.

[Hocevar, 1998] Hocevar, S.P. Deciding to adapt organizational architecture: Facilitators and inhibitors to change. *Proceedings for the 1998 Command & Control Research and Technology Symposium* (pp. 78-95). Monterey, CA, June 1998.

[Hocevar *et al.*, 1998] Hocevar, S.P., Kemple, W.G. & Benson, R. Translating simulation-based findings to an operational arena: Interpretations from research participant-experts. *Proceedings for the 1998 Command & Control Research and Technology Symposium* (pp. 54-65). Monterey, CA, June 1998.

[Hutchins *et al.*, 1999] Hutchins, S.G., Hocevar, S.P., & Kemple, W.G. Analysis of Team Communications in "Human-in-the-Loop" Experiments in Joint Command and Control. Paper presented at the 1999 Command and Control Research and Technology Symposium, Newport RI, June 1999.

[Kleinman & Young, 1997] Kleinman, D.L. & Young, P.W. *DDD-III User's guide, Version 2.1*, 1997.

[Kleinman *et al.*, 1996] Kleinman, D.L., Young, P.W. & G. Higgins. The DDD-III: A tool for empirical research in adaptive organizations. *Proceedings for the 1996 Symposium on Command and Control Research and Technology* (pp. 827-836). Monterey, CA, June 1996.

[Levchuk *et al.*, 1998] Levchuk, Y.N., Pattipati, K.R. & Kleinman, D.L. Designing adaptive organization to process a complex mission: Algorithms and applications. *Proceedings for the 1998 Command & Control Research and Technology Symposium* (pp. 11-32). Monterey, CA, June 1998.

[Mintzberg, 1993] Mintzberg, H. *Structure in fives: Designing effective organizations*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[Waller, 1999] Waller, M.J. The timing of adaptive group responses to nonroutine events. *Academy of Management Journal*, 42:127-137, 1999.

[Weick & Roberts, 1993] Weick, K.E. & Roberts, K.H. Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, 38: 357-381, 1993.